

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/47, 14/705, C12N 15/12	A1	(11) International Publication Number: WO 96/25432 (43) International Publication Date: 22 August 1996 (22.08.96)
---	-----------	---

(21) International Application Number: PCT/US95/01992

(22) International Filing Date: 17 February 1995 (17.02.95)

(71) Applicant (for all designated States except US): HUMAN
GENOME SCIENCES, INC. [US/US]; 9410 Key West
Avenue, Rockville, MD 20850-3338 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): LI, Yi [CN/US]; 16125
Howard Landing Drive, Gaithersburg, MD 20878 (US).
ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road,
Laytonsville, MD 20882 (US).

(74) Agents: OLSTEIN, Elliot, M. et al.; Carella, Byrne, Bain,
Gilfillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road,
Roseland, NJ 07068 (US).

(81) Designated States: AU, CA, CN, JP, KR, MX, NZ, US,
European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR,
IE, IT, LU, MC, NL, PT, SE).

Published

With international search report.

(54) Title: HUMAN G-PROTEIN COUPLED RECEPTOR

(57) Abstract

A human G-protein coupled receptor polypeptide and DNA (RNA) encoding such polypeptide and a procedure for producing such polypeptide by recombinant techniques is disclosed. Also disclosed are methods for utilizing such polypeptide for identifying antagonists and agonists to such polypeptide. Antagonists and agonists may be used therapeutically to inhibit or stimulate the G-protein coupled receptor. Also disclosed are assays for detecting mutations in the nucleic acid sequence encoding the G-protein coupled receptor.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

HUMAN G-PROTEIN COUPLED RECEPTOR

This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides. More particularly, the polypeptide of the present invention is a human 7-transmembrane receptor which has the greatest amino acid sequence homology to the human anaphylatoxin C5a receptor. The invention also relates to inhibiting the action of such polypeptides.

It is well established that many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers, e.g., cAMP (Lefkowitz, Nature, 351:353-354 (1991)). Herein these proteins are referred to as proteins participating in pathways with G-proteins or PPG proteins. Some examples of these proteins include the GPC receptors, such as those for adrenergic agents and dopamine (Kobilka, B.K., et al., PNAS, 84:46-50 (1987); Kobilka, B.K., et al., Science, 238:650-656 (1987); Bunzow, J.R., et al., Nature, 336:783-787 (1988)), G-proteins themselves, effector proteins, e.g., phospholipase C, adenylyl cyclase, and phosphodiesterase, and actuator proteins, e.g., protein

kinase A and protein kinase C (Simon, M.I., et al., Science, 252:802-8 (1991)).

For example, in one form of signal transduction, the effect of hormone binding is activation of an enzyme, adenylate cyclase, inside the cell. Enzyme activation by hormones is dependent on the presence of the nucleotide GTP, and GTP also influences hormone binding. A G-protein connects the hormone receptors to adenylate cyclase. G-protein was shown to exchange GTP for bound GDP when activated by hormone receptors. The GTP-carrying form then binds to an activated adenylate cyclase. Hydrolysis of GTP to GDP, catalyzed by the G-protein itself, returns the G-protein to its basal, inactive form. Thus, the G-protein serves a dual role, as an intermediate that relays the signal from receptor to effector, and as a clock that controls the duration of the signal.

The membrane protein gene superfamily of G-protein coupled receptors have been characterized as having seven putative transmembrane domains. The domains are believed to represent transmembrane α -helices connected by extracellular or cytoplasmic loops. G-protein coupled receptors include a wide range of biologically active receptors, such as hormone, viral, growth factor and neuro-receptors.

G-protein coupled receptors have been characterized as including these seven conserved hydrophobic stretches of about 20 to 30 amino acids, connecting at least eight divergent hydrophilic loops. Examples of G-protein family of coupled receptors includes dopamine receptors, calcitonin, adrenergic, endothelin, cAMP, adenosine, muscarinic, acetylcholine, serotonin, histamine, thrombin, kinin, follicle stimulating hormone, opsins and rhodopsins, odorant, cytomegalovirus receptors.

Most G-protein coupled receptors have single conserved cysteine residues in each of the first two extracellular loops which form disulfide bonds that are believed to

stabilize functional protein structure. The 7 transmembrane regions are designated as TM1, TM2, TM3, TM4, TM5, TM6, and TM7. TM3 is implicated in signal transduction.

Phosphorylation and lipidation (palmitylation or farnesylation) of cysteine residues can influence signal transduction of some G-protein coupled receptors. Most G-protein coupled receptors contain potential phosphorylation sites within the third cytoplasmic loop and/or the carboxy terminus. For several G-protein coupled receptors, such as the β -adrenoreceptor, phosphorylation by protein kinase A and/or specific receptor kinases mediates receptor desensitization.

The ligand binding sites of G-protein coupled receptors are believed to comprise a hydrophilic socket formed by several G-protein coupled receptors transmembrane domains, which socket is surrounded by hydrophobic residues of the G-protein coupled receptors. The hydrophilic side of each G-protein coupled receptor transmembrane helix is postulated to face inward and form the polar ligand binding site. TM3 has been implicated in several G-protein coupled receptors as having a ligand binding site, such as including the TM3 aspartate residue. Additionally, TM5 serines, a TM6 asparagine and TM6 or TM7 phenylalanines or tyrosines are also implicated in ligand binding.

G-protein coupled receptors can be intracellularly coupled by heterotrimeric G-proteins to various intracellular enzymes, ion channels and transporters (see, Johnson et al., Endoc., Rev., 10:317-331 (1989)). Different G-protein α -subunits preferentially stimulate particular effectors to modulate various biological functions in a cell. Phosphorylation of cytoplasmic residues of G-protein coupled receptors have been identified as an important mechanism for the regulation of G-protein coupling of some G-protein coupled receptors.

The anaphylatoxin C5a is a 74-amino acid polypeptide generated by cleavage of the alpha-chain of native C5 at a specific site by convertase of the blood complement system, as well as by enzymes of the coagulation system. In vivo, C5a is thought to play a significant role in the inflammatory response and in a number of clinical disorders (Goldstein, I.M., *Inflammation: Basic Principles and Clinical Correlates*, 309-323, Raven Press, New York (1988)). This peptide is a highly potent inflammatory agent, evoking dramatic responses in experimental animals (Bodammer, G. and Vogt, W., *Int. Arch. Allergy Appl. Immunol.*, 33:417-428 (1967)), and stimulating pulmonary, cardiac, vascular and gastrointestinal tissues in vitro (Stimler, N.P., et al., *Am. J. Pathol.*, 100:327-348 (1980)). C5a is a potent activator of polymorphonuclear neutrophils and macrophages, stimulating chemotaxis, hydrolytic enzyme release, and superoxide anion formation (Ward, P.A. and Newman, L.J., *J. Immunol.*, 102:93-99 (1969)).

Several reports have additionally demonstrated actions of this peptide on eosinophils, including chemotaxis and increased hexose uptake, in addition to its actions on mast cells and basophils (Hugli, T.E., *Biological Response Mediators and Modulators*, 99-116, Academic Press, New York (1983)). In addition, the anaphylatoxin has been shown to have a spasmogenic effect on various tissues; it stimulates smooth muscle contraction (Stimler, N.P., et al., *J. Immunol.*, 126:2258-2261 (1981)); induces histamine release from mast cells, promotes serotonin release from platelets (Meuer, S., et al., *J. Immunol.*, 126:1506-1509 (1981)), and increases vascular permeability (Jose, P.J., et al., *J. Immunol.*, 127:2376-2380 (1981)).

The interaction of C5a with polymorphonuclear leukocytes and other target cells and tissues results in increased histamine release, vascular permeability, smooth muscle contraction, and an influx into tissues of

inflammatory cells, including neutrophils, eosinophils and basophils (Hugli, T.E., Springer, Semin. Immunopathol., 7:193-219 (1981)). C5a may also play an important role in mediating inflammatory effects of phagocytic mononuclear cells that accumulate at sites of chronic inflammation (Allison, A.C., et al., H.U. Agents and Actions, 8:27 (1978)). C5a can induce chemotaxis in monocytes and cause them to release lysosomal enzymes in a manner analogous to the neutrophil responses elicited by these agents. C5a may have an immunoregulatory role by enhancing antibody, particularly at sites of inflammation (Morgan, E.L., et al., J. Exp. Med., 155:1412 (1982)).

A human C5a receptor cDNA clone has been isolated by expression cloning from a CDM8 expression library prepared from mRNA of Human myeloid HL-60 cells differentiated to the granulocyte phenotype with dibutyryladenine cyclic monophosphate (Boulay, F. et al., Biochemistry, 30:2993-2999 (1991)). Also, the human C5a receptor was cloned from U937 and HL-60 cells and identified by high affinity binding when expressed in COS-7 cells, (Gerard, N.P. and Gerard, C., Nature, 349:614-617 (1991)).

In accordance with one aspect of the present invention, there are provided novel G-protein coupled receptor polypeptides, as well as antisense analogs thereof and biologically active and diagnostically or therapeutically useful fragments and derivatives thereof. The polypeptides of the present invention are of human origin.

In accordance with another aspect of the present invention, there are provided isolated nucleic acid molecules encoding the human G-protein coupled receptor, including mRNAs, DNAs, cDNAs, genomic DNA as well as antisense analogs thereof and biologically active and diagnostically or therapeutically useful fragments thereof.

In accordance with a further aspect of the present invention, there is provided a process for producing such

polypeptides by recombinant techniques which comprises culturing recombinant prokaryotic and/or eukaryotic host cells, containing the human G-protein coupled receptor nucleic acid sequence, under conditions promoting expression of said protein and subsequent recovery of said protein.

In accordance with yet a further aspect of the present invention, there are provided antibodies against such polypeptides.

In accordance with another embodiment, there is provided a process for using the G-protein coupled receptors to screen for receptor antagonists and/or agonists and/or receptor ligands.

In accordance with still another embodiment of the present invention there is provided a process of using such agonists for stimulating the G-protein coupled receptor for the treatment of conditions related to the under-expression of the G-protein coupled receptors, for example, as a defense against bacterial infection, as a defense against viral infection, to stimulate the immunoregulatory effects of C5a, to treat immunodeficiency diseases and severe infections.

In accordance with another aspect of the present invention there is provided a process of using such antagonists for inhibiting the action of the G-protein coupled receptors for treating conditions associated with over-expression of the G-protein coupled receptors, for example, to treat asthma, bronchial allergy, chronic inflammation, systemic lupus erythematosus, vasculitis, rheumatoid arthritis, osteoarthritis, gout, certain auto-allergic diseases, transplant rejection, ulcerative colitis, in certain shock states, myocardial infarction, hypertension, abnormal cell growth and post-viral encephalopathies.

In accordance with another aspect of the present invention there are provided nucleic acid probes comprising nucleic acid molecules of sufficient length to specifically hybridize to human G-protein coupled receptor sequences.

In accordance with still another aspect of the present invention there are provided synthetic or recombinant G-protein coupled receptor polypeptides, conservative substitution and derivatives thereof, antibodies, anti-idiotypic antibodies, compositions and methods that can be useful as potential modulators of G-protein coupled receptor function, by binding to ligands or modulating ligand binding, due to their expected biological properties, which may be used in diagnostic, therapeutic and/or research applications.

In accordance with yet another object of the present invention, there is provided a diagnostic assay for detecting a disease or susceptibility to a disease related to a mutation in the G-protein coupled receptor nucleic acid sequence.

These and other aspects of the present invention should be apparent to those skilled in the art from the teachings herein.

The following drawings are illustrative of embodiments of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

Figure 1 shows the cDNA sequence and the corresponding deduced amino acid sequence of the putative mature G-protein coupled receptor of the present invention. The standard one-letter abbreviation for amino acids is used. Sequencing was performed using a 373 Automated DNA sequencer (Applied Biosystems, Inc.). Sequencing accuracy is predicted to be greater than 97% accurate.

Figure 2 illustrates an amino acid alignment of the G-protein coupled receptor of the present invention (top line) and a human C5a receptor (bottom line).

In accordance with an aspect of the present invention, there is provided an isolated nucleic acid (polynucleotide) which encodes for the mature polypeptide having the deduced amino acid sequence of Figure 1 (SEQ ID No. 2) or for the

mature polypeptide encoded by the cDNA of the clone deposited as ATCC Deposit No. 75982 on December 16, 1994.

A polynucleotide encoding a polypeptide of the present invention is predominantly expressed in peripheral lymphocytes. The polynucleotide of this invention was discovered in a cDNA library derived from a human activated neutrophil. It is structurally related to the G protein-coupled receptor family. It contains an open reading frame encoding a protein of 482 amino acid residues. The protein exhibits the highest degree of homology to a human C5a receptor with 26 % identity and 58 % similarity over the entire amino acid sequence.

While the G-protein coupled receptor has the highest degree of amino acid sequence homology to a human C5a receptor, there is also a significant degree of amino acid sequence homology to the human receptors for other ligands, for example N-formyl peptide, angiotensin, somatostatin, opioid, interleukin-8 (IL-8), bradykinin, thrombin and ATP receptors. Accordingly, while Applicant does not wish to limit the scientific theory underlying the present invention, the G-protein coupled receptor of the present invention may bind any one or a combination of the ligands identified above.

The polynucleotide of the present invention may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequence which encodes the mature polypeptide may be identical to the coding sequence shown in Figure 1 (SEQ ID No. 1) or that of the deposited clone or may be a different coding sequence which coding sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same mature polypeptide as the DNA of Figure 1 (SEQ ID No. 1) or the deposited cDNA.

The polynucleotide which encodes for the mature polypeptide of Figure 1 (SEQ ID No. 2) or for the mature polypeptide encoded by the deposited cDNA may include: only the coding sequence for the mature polypeptide; the coding sequence for the mature polypeptide and additional coding sequence such as a leader or secretory sequence or a proprotein sequence; the coding sequence for the mature polypeptide (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature polypeptide.

Thus, the term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention further relates to variants of the hereinabove described polynucleotides which encode for fragments, analogs and derivatives of the polypeptide having the deduced amino acid sequence of Figure 1 (SEQ ID No. 2) or the polypeptide encoded by the cDNA of the deposited clone. The variant of the polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-naturally occurring variant of the polynucleotide.

Thus, the present invention includes polynucleotides encoding the same mature polypeptide as shown in Figure 1 (SEQ ID No. 2) or the same mature polypeptide encoded by the cDNA of the deposited clone as well as variants of such polynucleotides which variants encode for a fragment, derivative or analog of the polypeptide of Figure 1 (SEQ ID No. 2) or the polypeptide encoded by the cDNA of the deposited clone. Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

As hereinabove indicated, the polynucleotide may have a coding sequence which is a naturally occurring allelic

variant of the coding sequence shown in Figure 1 (SEQ ID No. 1) or of the coding sequence of the deposited clone. As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded polypeptide.

The present invention also includes polynucleotides, wherein the coding sequence for the mature polypeptide may be fused in the same reading frame to a polynucleotide sequence which aids in expression and secretion of a polypeptide from a host cell, for example, a leader sequence which functions as a secretory sequence for controlling transport of a polypeptide from the cell. The polypeptide having a leader sequence is a preprotein and may have the leader sequence cleaved by the host cell to form the mature form of the polypeptide. The polynucleotides may also encode for a proprotein which is the mature protein plus additional 5' amino acid residues. A mature protein having a prosequence is a proprotein and is an inactive form of the protein. Once the prosequence is cleaved an active mature protein remains.

Thus, for example, the polynucleotide of the present invention may encode for a mature protein, or for a protein having a prosequence or for a protein having both a prosequence and a presequence (leader sequence).

The polynucleotides of the present invention may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptide of the present invention. The marker sequence may be a hexahistidine tag supplied by a pQE-9 vector to provide for purification of the mature polypeptide fused to the marker in the case of a bacterial host, or, for example, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to

an epitope derived from the influenza hemagglutinin protein (Wilson, I., et al., Cell, 37:767 (1984)).

The present invention further relates to polynucleotides which hybridize to the hereinabove-described sequences if there is at least 50% and preferably 70% identity between the sequences. The present invention particularly relates to polynucleotides which hybridize under stringent conditions to the hereinabove-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotides in a preferred embodiment encode polypeptides which either retain substantially the same biological function or activity as the mature polypeptide encoded by the cDNA of Figure 1 (SEQ ID No. 1) or the deposited cDNA, i.e. function as a G-protein coupled receptor or retain the ability to bind the ligand for the receptor even though the polypeptide does not function as a G-protein coupled receptor, for example, a soluble form of the receptor.

The deposit(s) referred to herein will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for purposes of Patent Procedure. These deposits are provided merely as convenience to those of skill in the art and are not an admission that a deposit is required under 35 U.S.C. §112. The sequence of the polynucleotides contained in the deposited materials, as well as the amino acid sequence of the polypeptides encoded thereby, are incorporated herein by reference and are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

The present invention further relates to a G-protein coupled receptor polypeptide which has the deduced amino acid sequence of Figure 1 (SEQ ID No. 2) or which has the amino acid sequence encoded by the deposited cDNA, as well as fragments, analogs and derivatives of such polypeptide.

The terms "fragment," "derivative" and "analog" when referring to the polypeptide of Figure 1 (SEQ ID No. 2) or that encoded by the deposited cDNA, means a polypeptide which either retains substantially the same biological function or activity as such polypeptide, i.e. functions as a G-protein coupled receptor, or retains the ability to bind the ligand or the receptor even though the polypeptide does not function as a G-protein coupled receptor, for example, a soluble form of the receptor. An analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide, preferably a recombinant polypeptide.

The fragment, derivative or analog of the polypeptide of Figure 1 (SEQ ID No. 2) or that encoded by the deposited cDNA may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed

to be within the scope of those skilled in the art from the teachings herein.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the G-protein coupled receptor genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The polynucleotides of the present invention may be employed for producing polypeptides by recombinant

techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the E. coli lac or trp, the phage lambda P_L promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in E. coli.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or

control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as E. coli, Streptomyces, Salmonella typhimurium; fungal cells, such as yeast; insect cells such as Drosophila S2 and Spodoptera Sf9; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are PKK232-8 and PCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, P_L and trp. Eukaryotic promoters include CMV immediate early, HSV

thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples including the SV40 enhancer on the late side of the

replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include E. coli, Bacillus subtilis, Salmonella typhimurium and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the

SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The G-protein coupled receptor polypeptides can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. Polypeptides of the invention may also include an initial methionine amino acid residue.

Fragments of the full length G-protein coupled receptor gene may be employed as a hybridization probe for a cDNA library to isolate the full length gene and to isolate other genes which have a high sequence similarity to the gene or similar biological activity. Probes of this type generally have at least 20 bases. Preferably, however, the probes have at least 30 bases and generally do not exceed 50 bases, although they may have a greater number of bases. The probe may also be used to identify a cDNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete G-protein coupled receptor gene including regulatory and promotor regions, exons, and

introns. As an example of a screen comprises isolating the coding region of the G-protein coupled receptor gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The G-protein coupled receptor of the present invention may be employed in a process for screening for antagonists and/or agonists for the receptor.

In general, such screening procedures involve providing appropriate cells which express the receptor on the surface thereof. In particular, a polynucleotide encoding the receptor of the present invention is employed to transfect cells to thereby express the G-protein coupled receptor. Such transfection may be accomplished by procedures as hereinabove described.

One such screening procedure involves the use of the melanophores which are transfected to express the G-protein coupled receptor of the present invention. Such a screening technique is described in PCT WO 92/01810 published February 6, 1992.

Thus, for example, such assay may be employed for screening for a receptor antagonist by contacting the melanophore cells which encode the G-protein coupled receptor with both the receptor ligand and a compound to be screened. Inhibition of the signal generated by the ligand indicates that a compound is a potential antagonist for the receptor, i.e., inhibits activation of the receptor.

The screen may be employed for determining an agonist by contacting such cells with compounds to be screened and determining whether such compound generates a signal, i.e., activates the receptor.

Other screening techniques include the use of cells which express the G-protein coupled receptor (for example, transfected CHO cells) in a system which measures extracellular pH changes caused by receptor activation, for example, as described in Science, volume 246, pages 181-296 (October 1989). For example, potential agonists or antagonists may be contacted with a cell which expresses the G-protein coupled receptor and a second messenger response, e.g. signal transduction or pH changes, may be measured to determine whether the potential agonist or antagonist is effective.

Another such screening technique involves introducing RNA encoding the G-protein coupled receptor into xenopus oocytes to transiently express the receptor. The receptor oocytes may then be contacted in the case of antagonist screening with the receptor ligand and a compound to be screened, followed by detection of inhibition of a calcium signal.

Another screening technique involves expressing the G-protein coupled receptor in which the receptor is linked to a phospholipase C or D. As representative examples of such cells, there may be mentioned endothelial cells, smooth muscle cells, embryonic kidney cells, etc. The screening for an antagonist or agonist may be accomplished as hereinabove described by detecting activation of the receptor or inhibition of activation of the receptor from the phospholipase second signal.

Another method involves screening for antagonists by determining inhibition of binding of labeled ligand to cells which have the receptor on the surface thereof. Such a method involves transfecting a eukaryotic cell with DNA encoding the G-protein coupled receptor such that the cell expresses the receptor on its surface and contacting the cell with a potential antagonist in the presence of a labeled form of a known ligand. The ligand can be labeled, e.g., by

radioactivity. The amount of labeled ligand bound to the receptors is measured, e.g., by measuring radioactivity of the receptors. If the potential antagonist binds to the receptor as determined by a reduction of labeled ligand which binds to the receptors, the binding of labeled ligand to the receptor is inhibited.

G-protein coupled receptors are ubiquitous in the mammalian host and are responsible for many normal and pathological biological functions. Accordingly, it is desirous to find compounds and drugs which stimulate the G-protein coupled receptors on the one hand and which can antagonize a G-protein coupled receptor on the other hand when it is desirable to inhibit the G-protein coupled receptor.

For example, agonists for G-protein coupled receptors may be employed for therapeutic purposes, such as the treatment of asthma, Parkinson's disease, acute heart failure, hypotension, urinary retention, and osteoporosis.

In general, antagonists to the G-protein coupled receptors may be employed for a variety of therapeutic purposes, for example, for the treatment of hypertension, angina pectoris, myocardial infarction, ulcers, asthma, allergies, benign prostatic hypertrophy and psychotic and neurological disorders, including schizophrenia, manic excitement, depression, delirium, dementia or severe mental retardation, dyskinesias, such as Huntington's disease or Gilles dila Tourett's syndrome, among others. G-protein coupled receptor antagonists have also been useful in reversing endogenous anorexia and in the control of bulimia.

Examples of G-protein coupled receptor antagonists include antibodies, or in some cases oligonucleotides, which bind to the G-protein coupled receptors but do not elicit a second messenger response such that the activity of the G-protein coupled receptors is prevented. Antibodies include anti-idiotypic antibodies which recognize unique determinants

generally associated with the antigen-binding site of an antibody. Potential antagonists also include proteins which are closely related to the ligand of the G-protein coupled receptors, i.e. a fragment of the ligand, which have lost biological function and when binding to the G-protein coupled receptors, elicit no response.

A potential antagonist also includes an antisense construct prepared through the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes for the mature polypeptides of the present invention, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (Triple helix -see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al, Science, 241:456 (1988); and Dervan et al., Science, 251: 1360 (1991)), thereby preventing transcription and the production of G-protein coupled receptors. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of mRNA molecules into G-protein coupled receptors (Antisense - Okano, J. Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of G-protein coupled receptors.

Another potential antagonist is a small molecule which binds to the G-protein coupled receptor, making it inaccessible to ligands such that normal biological activity is prevented. Examples of small molecules include but are not limited to small peptides or peptide-like molecules.

Potential antagonists also include a soluble form of a G-protein coupled receptor, e.g. a fragment of the receptors, which binds to the ligand and prevents the ligand from interacting with membrane bound G-protein coupled receptors.

Antagonists to the G-protein coupled receptor may also include a method of re-engineering the receptor such that the internal three-dimensional structure is maintained but the external structure is made hydrophilic.

The antagonists may be used generally as mediators of inflammatory responses, as immunoregulants and to treat all pathological conditions which result from anaphylaxis stimulated by the C5a polypeptide and mediated by the G-protein coupled receptor. These pathological conditions include asthma, bronchial allergy, chronic inflammation, systemic lupus erythematosus, vasculitis, serum sickness, angioedema, rheumatoid arthritis, osteoarthritis, gout, bullous skin diseases, hypersensitivity, pneumonitis, idiopathic pulmonary fibrosis, immune complex-mediated glomerulonephritis, psoriasis, allergic rhinitis, hypertension, adult respiratory distress syndrome, acute pulmonary disorders, endotoxin shock, hepatic cirrhosis, pancreatitis, inflammatory bowel diseases (including Crohn's disease and ulcerative colitis), thermal injury, gram-negative sepsis, necrosis in myocardial infarction, leukophoresis, exposure to medical devices (including, but not limited to, hemodialyzer membranes and extracorporeal blood circulation equipment), chronic hepatitis, transplant rejection, abnormal cell growth, for example tumors and cancers, post-viral encephalopathies, and/or ischemia induced myocardial or brain injury. These antagonist may also be used as prophylactics for such conditions as shock accompanying Dengue Hemorrhagic fever.

The agonists identified by the screening method as described above, may be employed to stimulate the G-protein coupled receptor to treat conditions related to an under-

expression of the receptor, which include defense against bacterial infection, stimulation of the immunoregulatory effects of C5a, immunodeficiency diseases, viral and other infections.

This invention additionally provides a method of treating abnormal conditions related to an excess of G-protein coupled receptor activity which comprises administering to a subject the antagonist as hereinabove described along with a pharmaceutically acceptable carrier in an amount effective to block binding of ligands to the G-protein coupled receptors and thereby alleviate the abnormal conditions.

The invention also provides a method of treating abnormal conditions related to an under-expression of G-protein coupled receptor activity which comprises administering to a subject a therapeutically effective amount of the agonist described above in combination with a pharmaceutically acceptable carrier, in an amount effective to enhance binding of ligands to the G-protein coupled receptor and thereby alleviate the abnormal conditions.

The soluble form of the G-protein coupled receptors, antagonists and agonists may be employed in combination with a suitable pharmaceutical carrier. Such compositions comprise a therapeutically effective amount of the antagonist or agonist, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The formulation should suit the mode of administration.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological

products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the pharmaceutical compositions may be employed in conjunction with other therapeutic compounds.

The pharmaceutical compositions may be administered in a convenient manner such as by the topical, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes. The pharmaceutical compositions are administered in an amount which is effective for treating and/or prophylaxis of the specific indication. In general, the pharmaceutical compositions will be administered in an amount of at least about 10 $\mu\text{g/kg}$ body weight and in most cases they will be administered in an amount not in excess of about 8 mg/Kg body weight per day. In most cases, the dosage is from about 10 $\mu\text{g/kg}$ to about 1 mg/kg body weight daily, taking into account the routes of administration, symptoms, etc.

The G-protein coupled receptor polypeptides and antagonists or agonists which are polypeptides, may be employed in accordance with the present invention by expression of such polypeptides *in vivo*, which is often referred to as "gene therapy."

Thus, for example, cells from a patient may be engineered with a polynucleotide (DNA or RNA) encoding a polypeptide *ex vivo*, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, cells may be engineered by procedures known in the art by use of a retroviral particle containing RNA encoding a polypeptide of the present invention.

Similarly, cells may be engineered *in vivo* for expression of a polypeptide *in vivo* by, for example, procedures known in the art. As known in the art, a producer cell for producing a retroviral particle containing RNA encoding the polypeptide of the present invention may be

administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a polypeptide of the present invention by such method should be apparent to those skilled in the art from the teachings of the present invention. For example, the expression vehicle for engineering cells may be other than a retrovirus, for example, an adenovirus which may be used to engineer cells *in vivo* after combination with a suitable delivery vehicle.

The invention also provides a method for determining whether a ligand not known to be capable of binding to the G-protein coupled receptor can bind to such receptor which comprises contacting a mammalian cell which expresses a G-protein coupled receptor with the ligand under conditions permitting binding of ligands to the G-protein coupled receptor, detecting the presence of a ligand which binds to the receptor and thereby determining whether the ligand binds to the G-protein coupled receptor. The systems hereinabove described for determining agonists and/or antagonists may also be employed for determining ligands which bind to the receptor.

This invention further provides a method of screening drugs to identify drugs which specifically interact with, and bind to, the human G-protein coupled receptors on the surface of a cell which comprises contacting a mammalian cell comprising an isolated DNA molecule encoding the G-protein coupled receptor with a plurality of drugs, determining those drugs which bind to the mammalian cell, and thereby identifying drugs which specifically interact with and bind to a human G-protein coupled receptor of the present invention.

This invention also provides a method of detecting expression of the G-protein coupled receptor on the surface of a cell by detecting the presence of mRNA coding for a G-protein coupled receptor which comprises obtaining total mRNA

from the cell and contacting the mRNA so obtained with a nucleic acid probe comprising a nucleic acid molecule of at least 15 nucleotides capable of specifically hybridizing with a sequence included within the sequence of a nucleic acid molecule encoding a human G-protein coupled receptor under hybridizing conditions, detecting the presence of mRNA hybridized to the probe, and thereby detecting the expression of the G-protein coupled receptor by the cell.

This invention is also related to the use of the G-protein coupled receptor genes as part of a diagnostic assay for detecting diseases or susceptibility to diseases related to the presence of mutations in the G-protein coupled receptor genes.

Individuals carrying mutations in the human G-protein coupled receptor genes may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a patient's cells, such as from blood, urine, saliva, tissue biopsy and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki et al., Nature, 324:163-166 (1986)) prior to analysis. RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid encoding the G-protein coupled receptor proteins can be used to identify and analyze G-protein coupled receptor mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled G-protein coupled receptor RNA or alternatively, radiolabeled G-protein coupled receptor antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic

mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, e.g., Myers et al., Science, 230:1242 (1985)).

Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g., Cotton et al., PNAS, USA, 85:4397-4401 (1985)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (e.g., Restriction Fragment Length Polymorphisms (RFLP)) and Southern blotting of genomic DNA.

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

The sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated region is used to

rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 500 or 600 bases; however, clones larger than 2,000 bp have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. FISH requires use of the clones from which the express sequence tag was derived, and the longer the better. For example, 2,000 bp is good, 4,000 is better, and more than 4,000 is probably not necessary to get good results a reasonable percentage of the time. For a review of this technique, see Verma et al., *Human Chromosomes: a Manual of Basic Techniques*, Pergamon Press, New York (1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, *Mendelian Inheritance in Man* (available on line through Johns Hopkins

University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, 1975, Nature, 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The present invention will be further described with reference to the following examples; however, it is to be understood that the present invention is not limited to such examples. All parts or amounts, unless otherwise specified, are by weight.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction

conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μ g of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 μ l of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μ g of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37°C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

Size separation of the cleaved fragments is performed using 8 percent polyacrylamide gel described by Goeddel, D. et al., *Nucleic Acids Res.*, 8:4057 (1980).

"Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T., et al., *Id.*, p. 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units to T4 DNA ligase ("ligase") per 0.5 μ g of approximately equimolar amounts of the DNA fragments to be ligated.

Unless otherwise stated, transformation was performed as described in the method of Graham, F. and Van der Eb, A., *Virology*, 52:456-457 (1973).

Example 1Bacterial Expression and Purification of G-protein coupled Receptor

The DNA sequence encoding the G-protein coupled receptor, ATCC # 75982, is initially amplified using PCR oligonucleotide primers corresponding to the 5' and 3' end sequences of the processed G-protein coupled receptor nucleic acid sequence (minus the signal peptide sequence) and the vector sequences 3' to the gene. Additional nucleotides corresponding to the G-protein coupled receptor nucleotide sequence are added to the 5' and 3' sequences respectively. The 5' oligonucleotide primer has the sequence 5' GACTAAAGCTTATGGCGTCTTTCTCTGCTGAG 3' (SEQ ID No. 3) contains a HindIII restriction enzyme site followed by 18 nucleotides of G-protein coupled receptor coding sequence starting from the presumed terminal amino acid of the processed protein codon. The 3' sequence 5' GAACTTCTAGACTTCACACAGTTGTACTATTT 3' (SEQ ID No. 4) contains complementary sequences to an XbaI site and is followed by 21 nucleotides of the gene. The restriction enzyme sites correspond to the restriction enzyme sites on the bacterial expression vector pQE-9 (Qiagen, Inc. Chatsworth, CA). pQE-9 encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter operator (P/O), a ribosome binding site (RBS), a 6-His tag and restriction enzyme sites. pQE-9 is then digested with HindIII and XbaI. The amplified sequences are ligated into pQE-9 and are inserted in frame with the sequence encoding for the histidine tag and the RBS. The ligation mixture is then used to transform E. coli strain available from Qiagen under the trademark M15/rep 4 by the procedure described in Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, Cold Spring Laboratory Press, (1989). M15/rep4 contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their

ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis. Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG ("Isopropyl-B-D-thiogalacto pyranoside") is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression. Cells are grown an extra 3 to 4 hours. Cells are then harvested by centrifugation. The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl. After clarification, solubilized G-protein coupled receptor is purified from this solution by chromatography on a Nickel-Chelate column under conditions that allow for tight binding by proteins containing the 6-His tag (Hochuli, E. et al., J. Chromatography 411:177-184 (1984)). The G-protein coupled receptor is eluted from the column in 6 molar guanidine HCl pH 5.0 and for the purpose of renaturation adjusted to 3 molar guanidine HCl, 100mM sodium phosphate, 10 mmolar glutathione (reduced) and 2 mmolar glutathione (oxidized). After incubation in this solution for 12 hours the protein is dialyzed to 10 mmolar sodium phosphate.

Example 2

Expression of Recombinant G-protein coupled Receptor in COS cells

The expression of plasmid, pG-protein coupled receptor HA is derived from a vector pCDNAI/Amp (Invitrogen) containing: 1) SV40 origin of replication, 2) ampicillin resistance gene, 3) E.coli replication origin, 4) CMV promoter followed by a polylinker region, a SV40 intron and

polyadenylation site. A DNA fragment encoding the entire pG-protein coupled receptor protein and a HA tag fused in frame to its 3' end is cloned into the polylinker region of the vector, therefore, the recombinant protein expression is directed under the CMV promoter. The HA tag correspond to an epitope derived from the influenza hemagglutinin protein as previously described (I. Wilson, H. Niman, R. Heighen, A. Cherenson, M. Connolly, and R. Lerner, 1984, Cell 37, 767). The infusion of HA tag to the target protein allows easy detection of the recombinant protein with an antibody that recognizes the HA epitope.

The plasmid construction strategy is described as follows:

The DNA sequence encoding for the G-protein coupled receptor, ATCC # 75982, is constructed by PCR on the full-length gene cloned using two primers: the 5' primer 5' GTCCAAGCTTGCCACCATGGGTCTTTCTCTGCT 3' (SEQ ID No. 5) contains a HindIII site followed by 18 nucleotides of G-protein coupled receptor coding sequence starting from the initiation codon; the 3' sequence 5' CTAGCTCGAGTCAAGCGTAGTCTGGGACGTCGT ATGGGTAGCACACAGTTGTACTATT 3' (SEQ ID No. 6) contains complementary sequences to an XhoI site, translation stop codon, HA tag and the last 15 nucleotides of the G-protein coupled receptor coding sequence (not including the stop codon). Therefore, the PCR product contains a HindIII site, G-protein coupled receptor coding sequence followed by HA tag fused in frame, a translation termination stop codon next to the HA tag, and an XhoI site. The PCR amplified DNA fragment and the vector, pCDNAI/Amp, are digested with HindIII and XhoI restriction enzymes and ligated. The ligation mixture is transformed into E. coli strain SURE (Stratagene Cloning Systems, La Jolla, CA) the transformed culture is plated on ampicillin media plates and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct

fragment. For expression of the recombinant G-protein coupled receptor, COS cells are transfected with the expression vector by DEAE-DEXTRAN method (J. Sambrook, E. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Laboratory Press, (1989)). The expression of the G-protein coupled receptor HA protein is detected by radiolabelling and immunoprecipitation method (E. Harlow, D. Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, (1988)). Cells are labelled for 8 hours with ³⁵S-cysteine two days post transfection. Culture media are then collected and cells are lysed with detergent (RIPA buffer (150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50mM Tris, pH 7.5). (Wilson, I. et al., Id. 37:767 (1984)). Both cell lysate and culture media are precipitated with a HA specific monoclonal antibody. Proteins precipitated are analyzed on 15% SDS-PAGE gels.

Example 3

Cloning and expression of G-protein coupled receptor using the baculovirus expression system

The DNA sequence encoding the full length G-protein coupled receptor protein, ATCC # 75982, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene:

The 5' primer has the sequence 5' CGGGATCCCTCCATG GCGTCTTTCTCTGCT 3' (SEQ ID No. 7) and contains a BamHI restriction enzyme site (in bold) followed by 4 nucleotides resembling an efficient signal for the initiation of translation in eukaryotic cells (J. Mol. Biol. 1987, 196, 947-950, Kozak, M.), which is just behind the first 18 nucleotides of the gene (the initiation codon for translation "ATG" is underlined).

The 3' primer has the sequence 5' CGGGATCCCGCTCACACAGTTGTACTATT 3' (SEQ ID No. 8) and contains the cleavage site for the restriction endonuclease BamHI and

18 nucleotides complementary to the 3' non-translated sequence of the G-protein coupled receptor gene. The amplified sequences are isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment is then digested with the endonuclease BamHI and then isolated again on a 1% agarose gel. This fragment is designated F2.

The vector pRG1 (modification of pVL941 vector, discussed below) is used for the expression of the G-protein coupled receptor protein using the baculovirus expression system (for review see: Summers, M.D. and Smith, G.E. 1987, A manual of methods for baculovirus vectors and insect cell culture procedures, Texas Agricultural Experimental Station Bulletin No. 1555). This expression vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV) followed by the recognition sites for the restriction endonuclease BamHI. The polyadenylation site of the simian virus (SV)40 is used for efficient polyadenylation. For an easy selection of recombinant viruses the beta-galactosidase gene from E.coli is inserted in the same orientation as the polyhedrin promoter followed by the polyadenylation signal of the polyhedrin gene. The polyhedrin sequences are flanked at both sides by viral sequences for the cell-mediated homologous recombination of co-transfected wild-type viral DNA. Many other baculovirus vectors could be used in place of pRG1 such as pAc373, pVL941 and pAcIM1 (Luckow, V.A. and Summers, M.D., Virology, 170:31-39).

The plasmid is digested with the restriction enzymes BamHI and then dephosphorylated using calf intestinal phosphatase by procedures known in the art. The DNA is then isolated from a 1% agarose gel as described above. This vector DNA is designated V2.

Fragment F2 and the dephosphorylated plasmid V2 are ligated with T4 DNA ligase. E.coli HB101 cells are then

transformed and bacteria identified that contained the plasmid (pBacG-protein coupled receptor) with the G-protein coupled receptor gene using the enzyme BamHI. The sequence of the cloned fragment is confirmed by DNA sequencing.

5 μ g of the plasmid pBacG-protein coupled receptor is co-transfected with 1.0 μ g of a commercially available linearized baculovirus ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA.) using the lipofection method (Felgner et al. Proc. Natl. Acad. Sci. USA, 84:7413-7417 (1987)).

1 μ g of BaculoGold™ virus DNA and 5 μ g of the plasmid pBacG-protein coupled receptor are mixed in a sterile well of a microtiter plate containing 50 μ l of serum free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards 10 μ l Lipofectin plus 90 μ l Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added dropwise to the Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours the transfection solution is removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. The plate is put back into an incubator and cultivation continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay performed similar as described by Summers and Smith (supra). As a modification an agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used which allows an easy isolation of blue stained plaques. (A detailed description of a "plaque assay" can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10).

Four days after the serial dilution, the viruses are added to the cells and blue stained plaques are picked with the tip of an Eppendorf pipette. The agar containing the recombinant viruses is then resuspended in an Eppendorf tube containing 200 μ l of Grace's medium. The agar is removed by a brief centrifugation and the supernatant containing the recombinant baculoviruses is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then stored at 4°C.

Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-G-protein coupled receptor at a multiplicity of infection (MOI) of 2. Six hours later the medium is removed and replaced with SF900 II medium minus methionine and cysteine (Life Technologies Inc., Gaithersburg). 42 hours later 5 μ Ci of 35 S-methionine and 5 μ Ci 35 S cysteine (Amersham) are added. The cells are further incubated for 16 hours before they are harvested by centrifugation and the labelled proteins visualized by SDS-PAGE and autoradiography.

Example 4

Expression pattern of G-protein coupled receptor in human tissue

Northern blot analysis is carried out to examine the levels of expression of G-protein coupled receptor in human tissues. Total cellular RNA samples are isolated with RNazol™ B system (Biotech Laboratories, Inc. Houston, TX). About 10 μ g of total RNA isolated from each human tissue specified is separated on 1% agarose gel and blotted onto a nylon filter (Sambrook, Fritsch, and Maniatis, Molecular Cloning, Cold Spring Harbor Press, (1989)). The labeling reaction is done according to the Stratagene Prime-It kit with 50ng DNA fragment. The labeled DNA is purified with a Select-G-50 column. (5 Prime - 3 Prime, Inc. Boulder, CO).

The filter is then hybridized with radioactive labeled full length G-protein coupled receptor gene at 1,000,000 cpm/ml in 0.5 M NaPO₄, pH 7.4 and 7% SDS overnight at 65°C. After being washed twice at room temperature and twice at 60°C with 0.5 x SSC, 0.1% SDS, the filter is then exposed at -70°C overnight with an intensifying screen. The message RNA for G-protein coupled receptor is abundant in peripheral lymphocytes.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, within the scope of the appended claims, the invention may be practiced otherwise than as particularly described.

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: LI, ET AL.
- (ii) TITLE OF INVENTION: Human G-protein coupled
Receptor
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
- (B) STREET: 6 BECKER FARM ROAD
- (C) CITY: ROSELAND
- (D) STATE: NEW JERSEY
- (E) COUNTRY: USA
- (F) ZIP: 07068
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5 INCH DISKETTE
- (B) COMPUTER: IBM PS/2
- (C) OPERATING SYSTEM: MS-DOS
- (D) SOFTWARE: WORD PERFECT 5.1
- (vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE: Concurrently
- (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA

- (A) APPLICATION NUMBER:
- (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: FERRARO, GREGORY D.
- (B) REGISTRATION NUMBER: 36,134
- (C) REFERENCE/DOCKET NUMBER: 325800-271

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 201-994-1700
- (B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 2040 BASE PAIRS
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CACGAGGAGA ACAGAAGAAG AGAAAGCTCA GCAAATTTTC TTGCCATACT TCATGACTTC 60
ACTGTGGCTA AGTGTGGGGA CCAGACAGGA CTCGTGGAGA CATCCAGGTG CTGAAGCCTT 120
CAGCTACTGT CTCAGTTTTT TGAAGTTTAG CAATGGCGTC TTTCTCTGCT GAGACCAATT 180
CAACTGACCT ACTCTCACAG CCATGGAATG AGCCCCCAGT AATTCTCTCC ATGGTCATTC 240
TCAGCCTTAC TTTTCTACTG GGATTGCCAG GCAATGGGCT GGTGCTGTGG GTGGCTGGCC 300
TGAAGATGCA GCGGACAGTG AACACAATTT GGTTCCTCCA CCTCACCTTG GCGGACCTCC 360
TCTGCTGCCT CTCCTTGCCC TTCTCGCTGG CTCACTTGGC TCTCCAGGGA CAGTGGCCCT 420
ACGGCAGGTT CCTATGCAAG CTCATCCCTT CCATCATTGT CCTCAACATG TTTGCCAGTG 480
TCTTCTGCT TACTGCCATT AGCCTGGATC GCTGTCTTGT GGTATTCAAG CCAATCTGGT 540
GTCAGAATCA TCGCAATGTA GGGATGGCCT GCTCTATCTG TGGATGTATC TGGGTGGTGG 600
CTTGTGTGAT GTGCATTCTT GTGTTCTGTG ACCGGGAAAT CTTCACTACA GACAACCATA 660
ATAGATGTGG CTACAAATTT GGTCTCTCCA GCTCATTAGA TTATCCAGAC TTTTATGGAG 720
ATCCACTAGA AAACAGGTCT CTTGAAAACA TTGTTTCAGCC GCCTGGAGAA ATGAATGATA 780
GGTTAGATCC TTCCTCTTTC CAAACAAATG ATCATCCTTG GACAGTCCCC ACTGTCTTCC 840
AACCTCAAAC ATTTCAAAGA CCTTCTGCAG ATTCACTCCC TAGGGGTTCT GCTAGGTTAA 900
CAAGTCAAAA TCTGTATTCT AATGTATTTA AACCTGCTGA TGTGGTCTCA CCTAAAATCC 960
CCAGTGGGTT TCCTATTGAA GATCACGAAA CCAGCCCACT GGATAACTCT GATGCTTTTC 1020
TCTCTACTCA TTAAAGCTG TTCCCTAGCG CTTCTAGCAA TTCCTTCTAC GAGTCTGAGC 1080
TACCACAAGG TTTCCAGGAT TATTACAATT TAGGCCAATT CACAGATGAC GATCAAGTGC 1140
CAACACCCCT CGTGGCAATA ACGATCACTA GGCTAGTGGT GGGTTTCCTG CTGCCCTCTG 1200
TTATCATGAT AGCCTGTTAC AGCTTCATTG TCTTCCGAAT GCAAAGGGGC CGCTTCGCCA 1260
AGTCTCAGAG CAAAACCTTT CGAGTGGCCG TGGTGGTGGT GGCTGTCTTT CTGTCTGCT 1320
GGACTCCATA CCACATTTT GGAGTCCTGT CATTGCTTAC TGACCCAGAA ACTCCCTTGG 1380
GGAAAACCTCT GATGTCCTGG GATCATGTAT GCATTGCTCT AGCATCTGCC AATAGTTGCT 1440
TTAATCCCTT CCTTTATGCC CTCTGGGGA AAGATTTTAG GAAGAAAGCA AGGCAGTCCA 1500
TTCAGGGAAT TCTGGAGGCA GCCTTCAGTG AGGAGCTCAC ACGTTCCACC CACTGTCCCT 1560
CAACAATGT CATTTAGAA AGAAATAGTA CAACTGTGTG AAAATGTGGA GCAGCCAACA 1620
AGCAGGGGCT CTTAGGCAAT CACATAGTGA AAGTTTATAA GAGGATGAAG TGATATGGTG 1680
AGCAGCGGAC TTCAAAACT GTCAAAGAAT CAATCCAGCG GTTCTCAAAC GGTACACAGA 1740
CTATTGACAT CAGCATCACC TAGAACTTG TTAGAAATGC AAATTCTCAA GCCGCATCCC 1800

AGACTTGCTG AATCGGAATC TCTGGGGGTT GGGACCCAGC AAGGGCACTT AACAAACCCC 1860
 CGTTTCTGAT TAATGCTAAA TGTAAGAATC ATTGTAAACA TTAGTTCTAT TTCTATCCCA 1920
 AACTAAGCTA TGTGAAATAA GAGAAGCTAC TTTGTTTTTA AATGATGTTG AATATTTGTC 1980
 GATATTTCCA TCATTAAATT TTTCCTTAGC ATTGTCTAAG TCAAAAAAAAA AAAAAAAAAA 2040

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 482 AMINO ACIDS

(B) TYPE: AMINO ACID

(C) STRANDEDNESS:

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ala	Ser	Phe	Ser	Ala	Glu	Thr	Asn	Ser	Thr	Asp	Leu	Leu	Ser
						5				10				15
Gln	Pro	Trp	Asn	Glu	Pro	Pro	Val	Ile	Leu	Ser	Met	Val	Ile	Leu
						20				25				30
Ser	Leu	Thr	Phe	Leu	Leu	Gly	Leu	Pro	Gly	Asn	Gly	Leu	Val	Leu
						35				40				45
Trp	Val	Ala	Gly	Leu	Lys	Met	Gln	Arg	Thr	Val	Asn	Thr	Ile	Trp
						50				55				60
Phe	Leu	His	Leu	Thr	Leu	Ala	Asp	Leu	Leu	Cys	Cys	Leu	Ser	Leu
						65				70				75

Pro	Phe	Ser	Leu	Ala	His	Leu	Ala	Leu	Gln	Gly	Gln	Trp	Pro	Tyr
						80				85				90
Gly	Arg	Phe	Leu	Cys	Lys	Leu	Ile	Pro	Ser	Ile	Ile	Val	Leu	Asn
						95				100				105
Met	Phe	Ala	Ser	Val	Phe	Leu	Leu	Thr	Ala	Ile	Ser	Leu	Asp	Arg
						110				115				120
Cys	Leu	Val	Val	Phe	Lys	Pro	Ile	Trp	Cys	Gln	Asn	His	Arg	Asn
						125				130				135
Val	Gly	Met	Ala	Cys	Ser	Ile	Cys	Gly	Cys	Ile	Trp	Val	Val	Ala
						140				145				150
Cys	Val	Met	Cys	Ile	Pro	Val	Phe	Val	Tyr	Arg	Glu	Ile	Phe	Thr
						155				160				165
Thr	Asp	Asn	His	Asn	Arg	Cys	Gly	Tyr	Lys	Phe	Gly	Leu	Ser	Ser
						170				175				180
Ser	Leu	Asp	Tyr	Pro	Asp	Phe	Tyr	Gly	Asp	Pro	Leu	Glu	Asn	Arg
						185				190				195
Ser	Leu	Glu	Asn	Ile	Val	Gln	Pro	Pro	Gly	Glu	Met	Asn	Asp	Arg
						200				205				210
Leu	Asp	Pro	Ser	Ser	Phe	Gln	Thr	Asn	Asp	His	Pro	Trp	Thr	Val
						215				220				225
Pro	Thr	Val	Phe	Gln	Pro	Gln	Thr	Phe	Gln	Arg	Pro	Ser	Ala	Asp
						230				235				240
Ser	Leu	Pro	Arg	Gly	Ser	Ala	Arg	Leu	Thr	Ser	Gln	Asn	Leu	Tyr
						245				250				255
Ser	Asn	Val	Phe	Lys	Pro	Ala	Asp	Val	Val	Ser	Pro	Lys	Ile	Pro
						260				265				270
Ser	Gly	Phe	Pro	Ile	Glu	Asp	His	Glu	Thr	Ser	Pro	Leu	Asp	Asn

275	280	285
Ser Asp Ala Phe Leu Ser Thr His Leu Lys Leu Phe Pro Ser Ala		
290	295	300
Ser Ser Asn Ser Phe Tyr Glu Ser Glu Leu Pro Gln Gly Phe Gln		
305	310	315
Asp Tyr Tyr Asn Leu Gly Gln Phe Thr Asp Asp Asp Gln Val Pro		
320	325	330
Thr Pro Leu Val Ala Ile Thr Ile Thr Arg Leu Val Val Gly Phe		
335	340	345
Leu Leu Pro Ser Val Ile Met Ile Ala Cys Tyr Ser Phe Ile Val		
350	355	360
Phe Arg Met Gln Arg Gly Arg Phe Ala Lys Ser Gln Ser Lys Thr		
365	370	375
Phe Arg Val Ala Val Val Val Val Ala Val Phe Leu Val Cys Trp		
380	385	390
Thr Pro Tyr His Ile Phe Gly Val Leu Ser Leu Leu Thr Asp Pro		
395	400	405
Glu Thr Pro Leu Gly Lys Thr Leu Met Ser Trp Asp His Val Cys		
410	415	420
Ile Ala Leu Ala Ser Ala Asn Ser Cys Phe Asn Pro Phe Leu Tyr		
425	430	435
Ala Leu Leu Gly Lys Asp Phe Arg Lys Lys Ala Arg Gln Ser Ile		
440	445	450
Gln Gly Ile Leu Glu Ala Ala Phe Ser Glu Glu Leu Thr Arg Ser		
455	460	465
Thr His Cys Pro Ser Asn Asn Val Ile Ser Glu Arg Asn Ser Thr		
470	475	480

Thr Val

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 32 BASE PAIRS

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GACTAAAGCT TATGGCGTCT TTCTCTGCTG AG

32

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 32 BASE PAIRS

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GAACTTCTAG ACTTCACACA GTTGTACTAT TT

32

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 33 BASE PAIRS

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTCCAAGCTT GCCACCATGG GTCCTTCTCT GCT

33

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 58 BASE PAIRS

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTAGCTCGAG TCAAGCGTAG TCTGGGACGT CGTATGGGTA GCACACAGTT GTACTATT 58

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 30 BASE PAIRS
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGGATCCCT CCATGGCGTC TTTCTCTGCT

30

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 29 BASE PAIRS
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGGGATCCCG CTCACACAGT TGTACTATT

29

WHAT IS CLAIMED IS:

1. An isolated polynucleotide selected from the group consisting of:
 - (a) a polynucleotide encoding the polypeptide having the deduced amino acid sequence of SEQ ID No. 2 or a fragment, analog or derivative of said polypeptide;
 - (b) a polynucleotide encoding the polypeptide having the amino acid sequence encoded by the cDNA contained in ATCC Deposit No. 75982 or a fragment, analog or derivative of said polypeptide.
2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.
3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.
4. The polynucleotide of Claim 1 wherein the polynucleotide is genomic DNA.
5. The polynucleotide of Claim 2 wherein said polynucleotide a polypeptide having the deduced amino acid sequence of SEQ ID No. 2.
6. The polynucleotide of Claim 2 wherein said polynucleotide encodes the polypeptide encoded by the cDNA of ATCC Deposit No. 75982.
7. The polynucleotide of Claim 1 having the coding sequence as shown in SEQ ID No. 1.
8. The polynucleotide of Claim 2 having the coding sequence deposited as ATCC Deposit No. 75982.
9. A vector containing the DNA of Claim 2.

10. A host cell genetically engineered with the vector of Claim 9.
11. A process for producing a polypeptide comprising: expressing from the host cell of Claim 10 the polypeptide encoded by said DNA.
12. A process for producing cells capable of expressing a polypeptide comprising genetically engineering cells with the vector of Claim 9.
13. An isolated DNA hybridizable to the DNA of Claim 2 and encoding a polypeptide having G-protein coupled receptor activity.
14. A polypeptide selected from the group consisting of (i) a polypeptide having the deduced amino acid sequence of SEQ ID No. 2 and fragments, analogs and derivatives thereof and (ii) a polypeptide encoded by the cDNA of ATCC Deposit No. 75982 and fragments, analogs and derivatives of said polypeptide.
15. The polypeptide of Claim 14 wherein the polypeptide has the deduced amino acid sequence of SEQ ID No. 2.
16. An antibody against the polypeptide of claim 14.
17. A compound which activates the polypeptide of claim 14.
18. A compound which inhibits activation of the polypeptide of claim 14.
19. A method for the treatment of a patient having need of activation of a G-protein coupled receptor polypeptide

comprising: administering to the patient a therapeutically effective amount of the compound of Claim 17.

20. A method for the treatment of a patient having need to inhibit activation of the G-protein coupled receptor comprising: administering to the patient a therapeutically effective amount of the compound of Claim 18.

21. The polypeptide of Claim 14 wherein the polypeptide is a soluble fragment of the G-protein coupled receptor and is capable of binding a ligand for the receptor.

22. A process for identifying compounds effective as antagonists or agonists comprising:

contacting a cell which expresses the polypeptide of claim 14 on the surface thereof with a known receptor ligand and a compound to be screened; and

determining if the compound inhibits or enhances activation of the receptor.

23. A process for determining whether a ligand not known to be capable of binding to a G-protein coupled receptor can bind thereto comprising:

contacting a mammalian cell which expresses the polypeptide of claim 14 on the surface thereof with a potential ligand;

detecting the presence of the ligand which binds to the polypeptide; and

determining whether the ligand binds to the polypeptide.

24. A method for diagnosing a disease or a susceptibility to a disease comprising:

detecting a mutation in the nucleic acid sequence encoding the a polypeptide of claim 14 in a sample derived from a host.

FIG. 1A

10 CACGAGGAGAACAGAAAGAGAAAGCTCAGCAAAATTTCTTGCCATACTTCATGACTTC 50
 70 ACTGTGGCTAAGTGTGGGACCAAGACAGGACTCGTGGAGACATCCAGGTGCTGAAGCCTT 110
 130 CAGCTACTGTCTCAGTTTTGAAGTTAGCAATGGCGTCTTTCTCTGCTGAGACCAATT 170
 190 M A S F S A E T N S 230
 210 CAACTGACCTACTCTCACAGCCATGGAATGAGCCCCCAGTAATTCTCTCCATGGTCATTC 230
 250 T D L L S Q P W N E P P V I L S M V I L 290
 270 TCAGCCTTACTTTTACTGGGATGCGCAGGCAATGGGCTGGTGTGCTGGTGGTGGCC 290
 310 S L T F L L G L P G N G L V L W V A G L 350
 330 TGAAGATGCAGCGGACAGTGAACACAATTGGTTCCTCCACCTCACCTTGGCGGACCTCC 350
 370 K M Q R T V N T I W F L H L T L A D L L 410
 390 TCTGCTGCCTCTCCTTGCCCTTCTCGCTGGCTCACTTGGCTCTCCAGGACAGTGGCCCT 410
 430 C C L S L P F S L A H L A L Q G Q W P Y 470
 450 ACGGCAGGTTCTTATGCAAGCTCATCCCTCCATCATTTGTCCTCAACATGTTTGCCAGTG 470
 490 G R F L C K L I P S I I V L N M F A S V 530
 510 TCTTCCTGCTTACTGCCATTAGCCTGGATCGCTGTGTGCTGTATCAAGCCAATCTGGT 530
 530 F L L T A I S L D R C L V V F K P I W C 530

MATCH WITH FIG. 1B

2 / 6

FIG. 1B

MATCH WITH FIG. 1A

550 570 590
 GTCAGATCATCGCAATGTAGGATGGCCCTGCTCTATCTGTGGATGTATCTGGTGGTGG
 Q N H R N V G M A C S I C G C I W V V A
 610 630 650
 CTTGTGTGATGTCATTCCTGTGTTCGTGTACCGGGAATCTTCACTACAGACAACCAT
 C V M C I P V F V Y R E I F T T D N H N
 670 690 710
 ATAGATGTGGCTACAAATTTGGTCTCTCCAGCTCATTAGATTATCCAGACTTTTATGGAG
 R C G Y K F G L S S L D Y P D F Y G D
 730 750 770
 ATCCACTAGAAAACAGGTCTCTTTGAAACATGTTCAGCCGCTGGAGAAATGAATGATA
 P L E N R S L E N I V Q P P G E M N D R
 790 810 830
 GGTTAGATCCTTCTTCCAAACAATGATCATCCTTGGACAGTCCCCACTGTCTTCC
 L D P S S F Q T N D H P W T V P T V F Q
 850 870 890
 AACCTCAAAACATTTCAAGACCTTCTGCAGATTCACTCCCTAGGGTTCTGCTAGGTAA
 P Q T F Q R P S A D S L P R G S A R L T
 910 930 950
 CAAGTCAAAATCTGTATCTAATGTATTAAACCTGCTGATGTGGTCTCACCTAAATCC
 S Q N L Y S N V F K P A D V V S P K I P
 970 990 1010
 CCAGTGGGTTTCCCTATTGAAGATCAGGAACCCAGCCCACTGGATAACTCTGATGCTTTC
 S G F P I E D H E T S P L D N S D A F L
 1030 1050 1070

MATCH WITH FIG. 1C

3 / 6

MATCH WITH FIG. 1B

FIG. 1C

TCCTACTCATTTAAAGCTGTTCCCTAGCGCTTCTAGCAATTCCTTCTACGAGTCTGAGC
 S T H L K L F P S A S S N S F Y E S E L
 1090 1110 1130
 TACCACAAGGTTTCCAGGATTATACAAATTTAGGCCAATTACAGATGACGATCAAGTGC
 P Q G F Q D Y Y N L G Q F T D D Q V P
 1150 1170 1190
 CAACACCCCTCGTGGCAATAACGATCACTAGCTAGTGGTGGTTCCTGCTGCCCTCTG
 T P L V A I T I T R L V V G F L L P S V
 1210 1230 1250
 TTATCATGATAGCCTGTTACAGCTTCATTTCTTCCGAATGCAAGGGCGCTTCGCCA
 I M I A C Y S F I V F R M Q R G R F A K
 1270 1290 1310
 AGTCTCAGAGCAAAACCTTTCGAGTGGCGGTGGTGGTGGTCTTTCTTGTCTGCT
 S Q S K T F R V A V V V V A V F L V C W
 1330 1350 1370
 GGACTCCATACACATTTTGGAGTCCCTGTTCATTTGCTTACTGACCCAGAAACTCCCTTGG
 T P Y H I F G V L S L L T D P E T P L G
 1390 1410 1430
 GGAAACTCTGATGTCCTGGGATCATGTATGCAATTGCTCTAGCATCTGCCAATAGTTGCT
 K T L M S W D H V C I A L A S A N S C F
 1450 1470 1490
 TTAATCCCTTCCTTATGCCCTCTTGGGGAAGATTTTAGGAAGAAAGCAAGCAGTCCA
 N P F L Y A L L G K D F R K K A R Q S I
 1510 1530 1550

MATCH WITH FIG. 1D

FIG. 1D

MATCH WITH FIG. 1C

```

TTCAGGGAATTCTGGAGGCAGCCCTTCAGTGAGGAGCTCACACGTTCCACCCACTGTCCCT
Q G I L E A A F S E E L T R S T H C P S
1570 1590 1610
CAAACAATGTCAATTCAGAAAGAAATAGTACAACCTGTGTGAAAAATGTGGAGCAGCCAAACA
N N V I S E R N S T T V *
1630 1650 1670
AGCAGGGGCTCTTAGGCAATCACATAGTGAAAGTTTATAAGAGGATGAAGTGATATGGTG
1690 1710 1730
AGCAGCGGACTTCAAAACTGTCAAGAATCAATCCAGCGGTTCTCAAACGGTACACAGA
1750 1770 1790
CTATTGACATCAGCATCACCTAGAAACTTGTTAGAAATGCAAAATTCCTCAAGCCGCATCCC
1810 1830 1850
AGACTTGCTGAATCGGAATCTCTGGGGTTGGGACCCAGCAAGGGCACTTAACAAACCCC
1870 1890 1910
CGTTTCTGATTAAATGCTAAATGTAAGAATCATTTGTAAACATTAGTTCTATTCTATCCCA
1930 1950 1970
AACTAAGCTATGTGAAATAAGAGAAGCTACTTTGTGTTTTTAAATGATGTTGAATATTGTC
1990 2010 2030
GATATTCCATCATTAATTTTTTCTTAGCATTTGTCTAAGTCAAAAAAATAAAAAA

```

MATCH WITH FIG. 2B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/01992

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07K 14/47, 14/705; C12N 15/12
US CL : 435/6, 69.1, 252.3, 320.1; 530/350, 388.22; 536/23.5

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 69.1, 252.3, 320.1; 530/350, 388.22; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
NONE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NATURE, Vol. 349, issued 14 February 1991, N. P. Gerard et al., "The chemotactic receptor for human C5a anaphylatoxin", pages 614-617, see entire document.	1-24
A	Biochemical Journal, Vol. 288, issued 1992, J. J. Perret et al., "Cloning and functional expression of the canine anaphylatoxin C5a receptor", pages 911-917, see entire document.	1-24

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	* T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
* A		document defining the general state of the art which is not considered to be of particular relevance
* E		earlier document published on or after the international filing date
* L		document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
* O		document referring to an oral disclosure, use, exhibition or other means
* P		document published prior to the international filing date but later than the priority date claimed
	* X	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
	* Y	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
	* Z	document member of the same patent family

Date of the actual completion of the international search

19 MAY 1995

Date of mailing of the international search report

30.05.95

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JOHN D. ULM

Telephone No. (703) 308-0196